# Estimation of Relative Camera Positions for Uncalibrated Cameras

*Richard I. Hartley*

G.E. CRD, Schenectady, NY, 12301.

**Abstract.** This paper considers, the determination of internal camera parameters from two views of a point set in three dimensions. A non-iterative algorithm is given for determining the focal lengths of the two cameras, as well as their relative placement, assuming all other internal camera parameters to be known. It is shown that this is all the information that may be deduced from a set of image correspondences.

## 1 Introduction

A non-iterative algorithm to solve the problem of relative camera placement was given by Longuet-Higgins ([4]). However, Longuet-Higgins's solution made assumptions about the camera that may not be justified in practice. In particular, it is assumed implicitly in his paper that the focal length of each camera is known, as is the principal point (the point where the focal axis of the camera intersects the image plane). Whereas it is often a safe assumption that the principal point of an image is at the center pixel, the focal length of the camera is not easily deduced, and will generally be unknown for images of unknown origin. In this paper a non-iterative algorithm is given for finding the focal lengths of the two cameras along with their relative placement, as long as other internal parameters of the cameras are known. It follows from the derivation of the algorithm, as well as from counting degrees of freedom that this is all the information that may be deduced about camera parameters from a set of image correspondences.

In this paper, the term magnification will be used instead of focal length, since it includes the equivalent effect of image enlargement.

## 2 The 8-Point Algorithm

First, I will derive the 8-point algorithm of Longuet-Higgins in order to fix notation and to gain some insight into its properties. Alternative derivations were given in [4] and [5]. Since we are dealing with homogeneous coordinates, we are interested only in values determined up to scale. Consequently we introduce the notation $A \approx B$ (where A and B are vectors or matrices) to indicate equality up to multiplication by a scale factor. Image space coordinates will usually be given in homogeneous coordinates as $(u, v, w)^\top$.

### 2.1 Algorithm Derivation

We consider the case of two cameras, one which is situated at the origin of object space coordinates, and one which is displaced from it. The two cameras may be represented by the transformation that they perform translating points from object space into image space coordinates. The two transformations are assumed to be

$$(u, v, w)^\top = (x, y, z)^\top \tag{1}$$

and

$$(u', v', w')^\top = R\left((x, y, z)^\top - (t_x, t_y, t_z)^\top\right) \tag{2}$$

where $R$ is a rotation matrix, the vectors $(u, v, w)^\top$ and $(u', v', w')^\top$ are the homogeneous coordinates of the image points, and $(x, y, z)^\top$ and $(t_x, t_y, t_z)^\top$ are non-homogeneous object space coordinates. Writing $T = (t_x, t_y, t_z)^\top$, and using homogeneous coordinates in both object and image space, the above relations may be written in matrix form as

$$(u, v, w)^\top = (I \mid 0)(x, y, z, 1)^\top = P_1(x, y, z, 1)^\top \tag{3}$$

and

$$(u', v', w')^\top = (R \mid -RT)(x, y, z, 1)^\top = P_2(x, y, z, 1)^\top \tag{4}$$

where $(I \mid 0)$ and $(R \mid -RT)$ are $3 \times 4$ matrices divided into a $3 \times 3$ block and a $3 \times 1$ column and $I$ is the identity matrix.

Now, I will define a transformation between the 2-dimensional projective plane of image coordinates in image 1 and the pencil of epipolar lines in the second image. As is well known, given a point $(u, v, w)^\top$ in image 1, the corresponding point in image 2 must lie on a certain epipolar line, which is the image under $P_2$ of the set $\mathcal{L}$ of all points $(x, y, z, 1)^\top$ which map under $P_1$ to $(u, v, w)^\top$. To determine this line one may identify two points in $\mathcal{L}$, namely the camera origin $(0, 0, 0, 1)^\top$ and the point at infinity, $(u, v, w, 0)^\top$. The images of these two points under $P_2$ are $-RT$ and $R(u, v, w)^\top$ respectively and the line that passes through these two points is given in homogeneous coordinates by the cross product,

$$(p, q, r)^\top = RT \times R(u, v, w)^\top = R\left(T \times (u, v, w)^\top\right) \ . \tag{5}$$

Here $(p, q, r)^\top$ represents the line $pu' + qv' + rw' = 0$. Representing by $S$ the matrix

$$S = S_T = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} \tag{6}$$

equation (5) may be written as

$$(p, q, r)^\top = RS(u, v, w)^\top \ . \tag{7}$$

Since the point $(u', v', w')^\top$ corresponding to $(u, v, w)^\top$ must lie on the epipolar line, we have the important relation

$$(u', v', w')Q(u, v, w)^\top = 0 \tag{8}$$

where $Q = RS$. This relationship is due to Longuet-Higgins ([4]).

As is well known, given 8 correspondences or more, the matrix $Q$ may be computed by solving a (possibly overdetermined) set of linear equations. In order to compute the second camera transform, $P_2$, it is necessary to factor $Q$ into the product $RS$ of a rotation matrix and a skew-symmetric matrix. Longuet-Higgins ([4]) gives a rather involved, and apparently numerically somewhat unstable method of doing this. I will give an alternative method of factoring the $Q$ matrix based on the Singular Value Decomposition ([1]). The following result may be verified.

**Theorem 1.** *A $3 \times 3$ real matrix $Q$ can be factored as the product of a rotation matrix and a non-zero skew symmetric matrix if and only if $Q$ has two equal non-zero singular values and one singular value equal to 0.*

A proof is contained in [2]. This theorem allows us to give an easy method of factoring any matrix into a product $RS$, when possible.

**Theorem 2.** *Suppose the matrix $Q$ can be factored into a product $RS$ where $R$ is orthogonal and $S$ is skew-symmetric. Let the Singular Value Decomposition of $Q$ be $UDV^\top$ where $D = \mathrm{diag}(k, k, 0)$. Then up to a scale factor the factorization is one of the following:*

$$S \approx VZV^\top \ ; \ R \approx UEV^\top \ \text{or} \ UE^\top V^\top \ ; \ Q \approx RS \ .$$

*where*

$$E = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \ Z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ . \tag{9}$$

*Proof.* That the given factorization is valid is true by inspection. That these are the only solutions is implicit in the paper of Longuet-Higgins ([4]). □

It may be verified that $T$ (the translation vector) in Theorem 2 is equal to $V.(0, 0, 1)^\top$ since this ensures that $ST = 0$ as required by (6). Furthermore $||T|| = 1$, which is a convenient normalization suggested in [4]. As remarked by Longuet-Higgins, the correct solution to the camera placement problem may be chosen based on the requirement that the visible points be in front of both cameras ([4]). There are four possible rotation/translation pairs that must be considered based on the two possible choices of $R$ and two possible signs of $T$. Therefore, since $UEV^\top V(0, 0, 1)^\top = U(0, 0, 1)^\top$ the requisite camera matrix $P_2 = (R \mid -RT)$ is equal to $(UEV^\top \mid -U(0, 0, 1)^\top)$ or one of the obvious alternatives.

## 2.2 Numerical Considerations

In any practical application, the matrix $Q$ found will not factor exactly in the required manner because of inaccuracies of measurement. In this case, the requirement will be to find the matrix closest to $Q$ that does factor into a product $RS$. Using the sum of squares of matrix entries as a norm (Frobenius norm [1]), we wish to find the matrix $Q' = RS$ such that $||Q - Q'||$ is minimized. The following theorem shows that the factorization given in the previous theorem is numerically optimal.

**Theorem 3.** *Let $Q$ be any $3 \times 3$ matrix and $Q = UDV^\top$ be its Singular Value Decomposition in which $D = \mathrm{diag}(r, s, t)$ and $r \geq s \geq t$. Define the matrix $Q'$ by $Q' = UD'V^\top$ where $D' = \mathrm{diag}(k, k, 0)$ and $k = (r+s)/2$. Then $Q'$ is the matrix closest to $Q$ in Frobenius norm which satisfies the condition $Q' = RS$, where $R$ is a rotation and $S$ is skew-symmetric. Furthermore, the factorization is given up to sign and scale by $R \approx UEV^\top$ or $UE^\top V^\top$ and $S \approx VZV^\top$.*

This theorem is plausible given the norm-preserving property of orthogonal transformations. However, its proof is not entirely obvious and falls beyond the scope of this paper.

## 2.3 Algorithm Outline

The algorithm for computing relative camera locations for calibrated cameras is as follows.

1. Find $Q$ by solving a set of equations of the form (8).
2. Find the Singular Value Decomposition $Q = UDV^\top$ , where $D = \text{diag}(a, b, c)$ and $a \geq b \geq c$.
3. The transformation matrices for the two cameras are $P_1 = (I \mid 0)$ and $P_2$ equal to one of the four following matrices.

$$\begin{array}{l} (UEV^\top \mid U(0,0,1)^\top) \\ (UEV^\top \mid -U(0,0,1)^\top) \\ (UE^\top V^\top \mid U(0,0,1)^\top) \\ (UE^\top V^\top \mid -U(0,0,1)^\top) \end{array}$$

The choice between the four transformations for $P_2$ is determined by the requirement that the point locations (which may be computed once the cameras are known [4]) must lie in front of both cameras. Geometrically, the camera rotations represented by $UEV^\top$ and $UE^\top V^\top$ differ from each other by a rotation through 180 degrees about the line joining the two cameras. Given this fact, it may be verified geometrically that a single pixel-to-pixel correspondence is enough to eliminate all but one of the four alternative camera placements.

# 3 Uncalibrated Cameras

If the internal camera calibration is not known, then the problem of finding the camera parameters is more difficult. In general one would like to allow arbitrary non-singular matrices $K$ describing internal camera calibration and consider camera matrices of the general form $(KR \mid -KRT)$, that is, general $3 \times 4$ matrices. Because $K$ is multiplied by a rotation, $R$, it may be assumed that $K$ is upper triangular. Allowing for an arbitrary scale factor, there are 5 remaining independent entries in $K$ representing camera parameters. Other authors ([6]) have allowed four internal camera parameters, namely principal point offsets in two directions and different scale factors in two directions. If however different scaling is allowed in two directions not necessarily aligned with the direction of the image-space axes, then one more parameter is needed, making up the 5.

It is too much to hope that from a set of image point correspondences one could retrieve the full set of internal camera parameters for a pair of cameras as well as the relative external positioning of the cameras. Indeed if $\{x_i\}$ are a set of points visible in a pair of cameras with transform matrices $P_1$ and $P_2$, and $G$ is an arbitrary non-singular $4 \times 4$ matrix, then replacing each $x_i$ by $G^{-1}x_i$ and each camera $P_j$ with $P_jG$ preserves the object-point to image-space correspondences. As may be seen, the internal parameters of one of the cameras, $P_1$ say, may be chosen arbitrarily. The situation is not helped by adding more cameras. This is in contrast to the case of calibrated cameras in which a finite number of solutions are possible ([2]). The question remains, therefore, how much can be deduced about the internal camera parameters from a set of image correspondences.

For uncalibrated cameras, a matrix $Q$ can be defined, analogous to the matrix defined for calibrated cameras, and this matrix may be computed given matched point pairs, according to (8). It may be observed that however many pairs of matched points are given, as far as determining camera models is concerned, the matrix $Q$ encapsulates all the information available, except as to which points lie behind or in front of the cameras. As remarked above, the choice of the four possible relative camera placements may be determined using just one matched point pair – the rest may be thrown away once $Q$ has been computed. To justify this observation it may be verified that a pair of matching points $(u,v,w)^\top$ and $(u',v',w')^\top$ correspond to a possible placement of an object point if and only if $(u',v',w')Q(u,v,w)^\top = 0$. This means that the addition of match points beyond 8 does not add any further information except numerical stability. Now, $Q$ has only 7 degrees of freedom consisting of 9 matrix entries, less one for arbitrary scale and one for the condition that $\det(Q) = 0$. (Theorem 1 does not hold for uncalibrated cameras.) Therefore, the total number of camera parameters that may be extracted from a set of image-point correspondences does not exceed 7. As shown by Longuet-Higgins, the relative camera placements account for 5 of these (not 6, since scale is indeterminate), and this paper accounts for two more, the camera magnification factors. It is not possible to extract any further information from $Q$, or hence from a set of matched points.

### 3.1 Form of the $Q$-matrix

Let $K_1$ and $K_2$ be two matrices representing the internal camera transformations of the two cameras and let $P_1 = (K_1 \mid 0)$ and $P_2 = (K_2R \mid -K_2RT)$ be the two camera transforms. The task is to obtain $R$, $T$, $K_1$ and $K_2$ given a set of image-point correspondences. For the present, the matrices $K_1$ and $K_2$ will be assumed arbitrary.

As before, it is possible to determine the epipolar line corresponding to a point $(u,v,w)^\top$ in image 1. The two points that must lie on the epipolar line are the images under $P_2$ of the camera centre $(0,0,0,1)^\top$ of the first camera and the point at infinity $\begin{pmatrix} K_1^{-1}(u,v,w)^\top \\ 0 \end{pmatrix}$. Transform $P_2$ takes these two points to the points $-K_2RT$ and $K_2RK_1^{-1}(u,v,w)^\top$. The line through these points is given by the cross product

$$K_2RT \times K_2RK_1^{-1}(u,v,w)^\top \ . \tag{10}$$

If $K$ is a square matrix, we use the notation $K^*$ to represent the cofactor matrix of $K$, that is the matrix defined by $K^*_{ij} = (-1)^{i+j}\det(K^{(ij)})$ where$K^{(ij)}$ is the matrix derived from $K$ by removing the $i$-th row and $j$-th column. If $K$ is non-singular, then it is well known that $K^* = \det(K).(K^\top)^{-1}$. In other words, $K^* \approx (K^\top)^{-1}$. The cofactor matrix is related to cross products in the following way.

**Lemma 4.** *If $a$ and $b$ are 3-dimensional column vectors and $K$ is a $3 \times 3$ matrix, then $Ka \times Kb \approx K^*(a \times b)$.*

Using this fact it is easy to evaluate the cross product (10).

$$K_2RT \times K_2RK_1^{-1}(u,v,w)^\top \approx K_2^*RK_1^{*\,-1}(K_1T \times (u,v,w)^\top) \tag{11}$$

Now, writing $S = S_{K_1T}$ as defined in (6), we have a formula for the epipolar line corresponding to the point $(u,v,w)^\top$ in image 1 :

$$(p,q,r)^\top \approx K_2^*R{K_1}^\top S(u,v,w)^\top \ . \tag{12}$$

Furthermore, setting $Q = K_2^*R{K_1}^\top S$ we have the formula

$$(u',v',w')Q(u,v,w)^\top = 0 \ . \tag{13}$$

An alternative factorization for $Q$ that may be derived from (10) and Lemma 4 is

$$Q \approx (K_2^{-1})^\top RSK_1^{-1} \tag{14}$$

where $S = S_T$ as given by (6).

### 3.2 Factorization of $Q$

Our goal, given $Q$, is to find the factorization $Q \approx K_2^* R K_1{}^\top S$. As before, we use the Singular Value Decomposition, $Q = UDW^\top$. By multiplying by $-1$ if necessary, $U$ and $V$ may be chosen such the $\det(U) = \det(V) = +1$ so that $U^* = U$ and $V^* = V$. Since $Q$ is singular, the diagonal matrix $D$ equals $\mathrm{diag}(r, s, 0)$ where $r$ and $s$ are positive constants. Since $QW(0,0,1)^\top = 0$, it follows that $SW(0,0,1)^\top = 0$ since $K_2^* R K_1{}^\top$ is non-singular, and so $S \approx WZW^\top$ where $Z$ is given in (9). The general solution to the problem of factoring $Q$ into a product $R'S'$, where $R'$ is non-singular and $S'$ is skew-symmetric is therefore given by

$$Q = (UX_{\alpha,\beta,\gamma}E^\top W^\top).(WZW^\top) \tag{15}$$

where $X_{\alpha,\beta,\gamma}$ is given by

$$X_{\alpha,\beta,\gamma} = \begin{pmatrix} r & 0 & \alpha \\ 0 & s & \beta \\ 0 & 0 & \gamma \end{pmatrix} \tag{16}$$

and $\alpha$, $\beta$ and $\gamma$ are arbitrary constants. The two bracketed expressions are $R'$ and $S'$ respectively and the factorization is unique (except for the variables $\alpha$, $\beta$ and $\gamma$) up to scale. In contrast to the situation in Section 2.1 we do not need to consider the alternate solution in which $E^\top$ is replaced by $E$, since that is taken care of by the undetermined values $\alpha$, $\beta$ and $\gamma$. Since both $E$ and $W$ are orthogonal matrices, we write $V = WE$, and $V$ is also orthogonal.

Now, we turn our attention to the matrix $R' = UX_{\alpha,\beta,\gamma}V^\top$. For some values of $\alpha$, $\beta$ and $\gamma$, it must be true that $R' \approx K_2^* R K_1^{*\,-1}$, where $R$ is a rotation matrix. From this it follows that $R \approx K_2^{*\,-1} R' K_1^*$. We now apply the property that a rotation matrix is equal to its cofactor matrix, (inverse transpose). This means that $K_2^{*\,-1} R' K_1^* \approx K_2^{-1} R'^* K_1$ or

$$K_2 K_2{}^\top R' \approx R'^* K_1 K_1{}^\top \ . \tag{17}$$

Since $R' \approx UX_{\alpha,\beta,\gamma}V^\top$ , it follows that $R'^* \approx UX^*_{\alpha,\beta,\gamma}V^\top$ where $X^*_{\alpha,\beta,\gamma}$ is the matrix

$$X^*_{\alpha,\beta,\gamma} = \begin{pmatrix} s\gamma & 0 & 0 \\ 0 & r\gamma & 0 \\ -s\alpha & -r\beta & rs \end{pmatrix} \tag{18}$$

and so from (17)

$$(K_2 K_2{}^\top) U X_{\alpha,\beta,\gamma} V^\top \approx U X^*_{\alpha,\beta,\gamma} V^\top (K_1 K_1{}^\top) \ . \tag{19}$$

At this point, it is necessary to specialize to the case where $K_1$ and $K_2$ are of the simple form $K_1 = \mathrm{diag}(1,1,k_1)$ and $K_2 = \mathrm{diag}(1,1,k_2)$. In this case, $k_1$ and $k_2$ are the inverses of the magnification factors. If the entries of $UX_{\alpha,\beta,\gamma}V^\top$ are $(f_{ij})$ and those of $UX^*_{\alpha,\beta,\gamma}V^\top$ are $(g_{ij})$, then multiplying by $(K_2K_2{}^\top)$ and $(K_1K_1{}^\top)$ respectively gives an equation

$$\begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ k_2^2 f_{31} & k_2^2 f_{32} & k_2^2 f_{33} \end{pmatrix} = x \begin{pmatrix} g_{11} & g_{12} & k_1^2 g_{13} \\ g_{21} & g_{22} & k_1^2 g_{23} \\ g_{31} & g_{32} & k_1^2 g_{33} \end{pmatrix} \tag{20}$$

where the $f_{ij}$ and $g_{ij}$ linear expressions in $\alpha$, $\beta$ and $\gamma$, and $x$ is an unknown scale factor. The top left hand block of (20) comprises a set of equations of the form

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = x \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \ . \tag{21}$$

If the scale factor were known, then this system could be solved for $\alpha$, $\beta$ and $\gamma$ as a set of linear equations. Unfortunately, $x$ is not known, and it is necessary to find the value of $x$ before solving the set of linear equations. Since the entries of the matrices on both sides of (21) are linear expressions in $\alpha$, $\beta$ and $\gamma$, it is possible to rewrite (21) in the form

$$M_1(\alpha, \beta, \gamma, 1)^\top - x\, M_x(\alpha, \beta, \gamma, 1)^\top = 0 \ , \tag{22}$$

where $M_1$ and $M_x$ are $4 \times 4$ matrices, each row of $M_1$ or $M_x$ corresponding to one of the four entries in the matrices in (21). Such a set of equations has a solution only if $\det(M_1 - x\, M_x) = 0$. This leads to a polynomial equation of degree 4 in $x$ : $p(x) = \det(M_1 - x\, M_x) = 0$. It will be seen later that this polynomial reduces to a quadratic.

The form of the matrix $M_1$ may be written out explicitly. Let $X_{\alpha,\beta,\gamma}$ be written in the form $\alpha.\Delta_{13} + \beta.\Delta_{23} + \gamma.\Delta_{33} + (r.\Delta_{11} + s\Delta_{22})$, where $\Delta_{ij}$ is the matrix having a one in position $i, j$ and zeros elsewhere. Then,

$$UX_{\alpha,\beta,\gamma}V^\top = \alpha U\Delta_{13}V^\top + \beta U\Delta_{23}V^\top + \gamma U\Delta_{33}V^\top + rU\Delta_{11}V^\top + sU\Delta_{22}V^\top \ .$$

It may be verified that the the $p, q$-th entry of the matrix $U\Delta_{ij}V^\top$ is equal to $U_{pi}V_{qj}$. Now, suppose that the rows of $M_1$ are ordered corresponding to the entries $f_{11}, f_{12}, f_{21}$ and $f_{22}$ of $UX_{\alpha,\beta,\gamma}V^\top$. Then

$$\begin{pmatrix} f_{11} \\ f_{12} \\ f_{21} \\ f_{22} \end{pmatrix} = \begin{pmatrix} U_{11}V_{13} & U_{12}V_{13} & U_{13}V_{13} & r.U_{11}V_{11}+s.U_{12}V_{12} \\ U_{11}V_{23} & U_{12}V_{23} & U_{13}V_{23} & r.U_{11}V_{21}+s.U_{12}V_{22} \\ U_{21}V_{13} & U_{22}V_{13} & U_{23}V_{13} & r.U_{21}V_{11}+s.U_{22}V_{12} \\ U_{21}V_{23} & U_{22}V_{23} & U_{23}V_{23} & r.U_{21}V_{21}+s.U_{22}V_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ 1 \end{pmatrix} \tag{23}$$

and $M_1$ is the matrix in this expression. The exact form of the matrix $M_x$ may be computed in a similar manner.

$$M_x = \begin{pmatrix} -s.U_{13}V_{11} & -r.U_{13}V_{12} & r.U_{12}V_{12}+s.U_{11}V_{11} & rs.U_{13}V_{13} \\ -s.U_{13}V_{21} & -r.U_{13}V_{22} & r.U_{12}V_{22}+s.U_{11}V_{21} & rs.U_{13}V_{23} \\ -s.U_{23}V_{11} & -r.U_{23}V_{12} & r.U_{22}V_{12}+s.U_{21}V_{11} & rs.U_{23}V_{13} \\ -s.U_{23}V_{21} & -r.U_{23}V_{22} & r.U_{22}V_{22}+s.U_{21}V_{21} & rs.U_{23}V_{23} \end{pmatrix} \tag{24}$$

With the help of a symbolic algebraic manipulation program such as Mathematica ([7]) three identities may easily be established by direct computation :

$$\det(M_x) = 0 \ , \ \det(M_1) = 0 \ , \ \det(M_1 + M_x) + det(M_1 - M_x) = 0 \ .$$

From this it follows easily that $p(x) = \det(M_1 - x\, M_x) = a_1 x + a_3 x^3$. The root $x = 0$ of this polynomial may safely be ignored, since according to (21) it would imply that $f_{ij} = 0$ for $i, j \leq 2$, and hence that $R$ is singular, which by assumption it is not. Thus $p(x)$ reduces to a quadratic as promised, and this quadratic has two roots of equal magnitude and opposite sign. It is possible that $p(x)$ has no real root, which indicates that no real solution is possible given the assumed camera model. This may mean that the position of the principal points have been wrongly guessed. For a different value of each principal point (that is, a translation of image space coordinates) a solution may be possible, but the solution will be dependent on the particular translations chosen.

Supposing, however, that $x$ is a real root of $p(x)$, the values of $\alpha$, $\beta$ and $\gamma$ may be determined by solving the set of equations given in (21). Finally, the values of $k_1$ and $k_2$ may be read off from equation (20). In particular,

$$\begin{array}{llll} k_2^2 &= x.g_{31}/f_{31} &= x.g_{32}/f_{32} & \text{(i)} \\ k_1^2 &= f_{13}/x.g_{13} &= f_{23}/x.g_{23} & \text{(ii)} \\ k_2^2 f_{33} &= x.k_1^2 g_{33} \ . & & \text{(iii)} \end{array} \tag{25}$$

The apparent redundancy in the equations (25) is resolved by the following proposition.

**Proposition 5.**
1. *If $x$ is either of the roots of $p(x)$, then the two expressions $x\,g_{31}/f_{31}$ and $x\,g_{32}/f_{32}$ for $k_1^2$ in (25.i) are equal. Similarly, the two expressions for $k_2^2$ in (25.ii) are equal and the relationship (25.iii) is always true.*
2. *Values $k_1^2$ and $k_2^2$ are either both positive or both negative.*
3. *The estimated values of $k_1^2$ corresponding to the two opposite roots of $p(x)$ are the same. The same holds for the two values of $k_2^2$.*

Proof of this proposition is beyond the scope of this paper. The case where $k_1^2$ and $k_2^2$ are negative implies as before that no solution is possible. Once again, selecting a different value for the principal points (origin of image-space coordinates) may lead to a solution.

At this point, it is possible to continue and compute the values of the rotation matrix directly. However, it turns out to be more convenient, now that the values of the magnification are known, to revert to the case of a calibrated camera. More particularly, we observe that according to (14), $Q$ may be written as $Q = K_2^{-1}Q'K_1^{-1}$ where $Q' = RS$, and $R$ is a rotation matrix. The original method of Section 2.3 may now be used to solve for the camera matrices derived from $Q'$. In this way, we find camera models $P_1 = (I \mid 0)$ and $P_2 = (R \mid -RT)$ for the two cameras corresponding to $Q'$. Taking account of the magnification matrices $K_1$ and $K_2$, the final estimates of the camera matrices are $(K_1 \mid 0)$ and $(K_2R \mid -K_2RT)$.

In practice it has been observed that greater numerical accuracy is obtained by repeating the computation of $k_1$ and $k_2$ after replacing $Q$ by $Q'$. The values of $k_1$ and $k_2$ computed from $Q'$ are very close to 1 and may be used to revise the computed magnifications very slightly. However, such a revision is necessary only because of numerical round-off error in the algorithm and is not strictly necessary.

### 3.3 Algorithm Outline

Although the mathematical derivation of this algorithm is at times complex, the implementation is not particularly difficult. The steps of the algorithm are reiterated here.

1. Compute a matrix $Q$ such that $(u_i', v_i', 1)^\top Q(u_i, v_i, 1) = 0$ for each of several matched pairs (at least 8 in number) by a linear least-squares method.
2. Compute the Singular Value Decomposition $Q \approx UDW^\top$ with $\det(U) = \det(V) = +1$ and set $r$ and $s$ to equal the two largest singular values. Set $V = WE$.
3. Form the matrices $M_1$ and $M_x$ given by (23) and (24) and compute the determinant $p(x) = \det(M_1 - x\,M_x) = a_1x + a_3x^3$.
4. If $-a_1/a_3 < 0$ no solution is possible, so stop. Otherwise, let $x = \sqrt{-a_1/a_3}$, one of the roots of $p(x)$.
5. Solve the equation $(M_1 - x\,M_x)(\alpha, \beta, \gamma, 1)^\top = 0$ to find $\alpha$, $\beta$ and $\gamma$ and use these values to form the matrices $X_{\alpha,\beta,\gamma}$ and $X^*_{\alpha,\beta,\gamma}$ given by (16) and (18).
6. Form the products $UX_{\alpha,\beta,\gamma}V^\top$ and $UX^*_{\alpha,\beta,\gamma}V^\top$ and observe that the four top left elements of these matrices are the same.
7. Compute $k_1$ and $k_2$ from the equations (25) where $(f_{ij})$ and $(g_{ij})$ are the entries of the matrices $UX_{\alpha,\beta,\gamma}V^\top$ and $UX^*_{\alpha,\beta,\gamma}V^\top$ respectively. If $k_1$ and $k_2$ are imaginary, then no solution is possible, so stop.
8. Compute the matrix $Q' = K_2QK_1$ where $K_1$ and $K_2$ are the matrices $\mathrm{diag}(1, 1, k_1)$ and $\mathrm{diag}(1, 1, k_2)$ respectively.

9. Compute the Singular Value Decomposition of $Q' = U'D'V'^{\top}$.
10. Set $P_1 = (K_1 \mid 0)$ and set $P_2$ to be one of the matrices

$$\begin{array}{l} (K_2U'EV'^{\top} \mid K_2U'(0,0,1)^{\top}) \\ (K_2U'E^{\top}V'^{\top} \mid K_2U'(0,0,1)^{\top}) \\ (K_2U'EV'^{\top} \mid -K_2U'(0,0,1)^{\top}) \\ (K_2U'E^{\top}V'^{\top} \mid -K_2U'(0,0,1)^{\top}) \end{array}$$

according to the requirement that the matched points must lie in front of both cameras.

# 4 Practical Results

This algorithm has been encoded in C and tested on a variety of examples. In the first test, a set of 25 matched points was computed synthetically, corresponding to an oblique placement of two cameras with equal magnification values of 1003. The principal point offset was assumed known. The solution to the relative camera placement problem was computed. The two cameras were computed to have magnifications of 1003.52 and 1003.71, very close to the original. Camera placements and point positions were computed and were found to match the input pixel position data within limits of accuracy. Similarly, the positions in 3-space of the object points matched the known positions to within one part in $10^4$.

The algorithm was also tested out on a set of matched points derived from a stereo-matching program, STEREOSYS ([3]). A set of 124 matched points were found by an unconstrained hierarchical search. The two images used were $1024 \times 1024$ aerial overhead images of the Malibu region with about 40% overlap. The algorithm described here was applied to the set of 124 matched points and relative camera placements and object-point positions were computed. The computed model was then evaluated against the original data. Consequently, the computed camera models were applied to the computed 3-D object points to give new pixel locations which were then compared with the original reference pixel data. The RMS pixel error was found to be 0.11 pixels. In other words, the derived model matches the actual data with a standard deviation of 0.11 pixels. This shows the accuracy not only of the derived camera model, but also the accuracy of the point-matching algorithms.

# References


1. Atkinson, K. E., "An Introduction to Numerical Analysis," John Wiley and Sons, New York, Second Edition 1989.
2. Faugeras, O. and Maybank, S., "Motion from Point Matches : Multiplicity of Solutions," International Journal of Computer Vision, **4**, (1990), 225–246.
3. Hannah, M.J., "Bootstrap Stereo," Proc. Image Understanding Workshop, College Park, MD, April 1980, 201–208.
4. Longuet-Higgins, H. C., "A computer algorithm for reconstructing a scene from two projections," Nature, Vol. 293, **10**, Sept. 1981.
5. Porrill, J and Pollard, S., "Curve Fitting and Stereo Calibration," Proceedings of the British Machine Vision Conference, University of Oxford, Sept 1990, pp 37–42.
6. Strat, T. M., "Recovering the Camera Parameters from a Transformation Matrix," Readings in Computer Vision, pp 93-100, Morgan Kauffamann Publishers, Inc, Los Altos, Ca., 1987.
7. Wolfram, S., Mathematica, "A System for Doing Mathematics by Computer," Addison-Wesley, Redwood City, California, 1988.


This article was processed using the LaTeX macro package with ECCV92 style